Theory and Practice of Data Assimilation in Ocean Modeling

Robert N. Miller
College of Oceanic and Atmospheric Sciences
Oregon State University
Oceanography Admin. Bldg. 104
Corvallis, OR 97331-5503
phone: (541) 737-4555 fax: (541) 737-2064 email: miller@oce.orst.edu
Award Number N00014-92-J-1595
http://tangaroa.oce.orst.edu/group.html

LONG-TERM GOAL

The long range goal of this project is to combine computational models with observational data to form the best picture of the ocean as an evolving system, and use that picture to understand the physical influences which govern the ocean's behavior. Oceanic observations are sparse and models are limited in accuracy, but taken together, one can form a quantitative description of the state of the ocean that is superior to any based on either models or data alone. Along with the goals of analysis and prediction, we seek reliable estimates of the errors in our results. We expect our results to have implications beyond data assimilation. In particular, we hope this research will lead to enhanced understanding of the implications of nonlinearity and randomness for predictability of the ocean and atmosphere.

It is our long range goal to develop efficient methods for estimating useful statistical measures of errors in stochastic forecast models. Since the probability density function (PDF's) of nonlinear stochastic models are not, in general, Gaussian, we must find methods for forecast evaluation based information about the particular PDF generated by the model.

OBJECTIVES

The principle objective of this project is the development, implementation and validation of practical data assimilation methods for synoptic ocean models. By ``data assimilation'' we mean the construction of a composite estimate of the state of the ocean based on a combination of observed data with computational model output. Since data assimilation methods which give the most and best information are highly resource intensive, and often not practical for use with detailed models, we are particularly interested in the price paid in terms of accuracy and confidence for using economical but suboptimal data assimilation methods.

Direct calculation of full PDF's is not feasible for practical models of the ocean or atmosphere, but useful approximations to the PDF can be calculated from Monte-Carlo experiments, by virtue of the fact that the number of truly independent degrees of freedom in practical models is very much smaller than the dimension of the state vector. This intuition is the motivation for the ensemble methods that have become popular in recent years.

Our experience with Monte-Carlo methods in simplified systems has led us to investigate the details of methods for ensemble generation that have been presented in the community. The motivation for these specialized methods for generating ensembles is precisely the specification of the PDF of a complex model whose behavior is believed to be captured by a relatively small number of independent degrees of freedom. By detailed study of the behavior of ensembles in increasingly complex models, we hope to gain the insights necessary to generate the most efficient ensembles, which should in turn lead to the error estimates necessary for data assimilation systems and prior estimates of forecast accuracy.

Optimized methods require accurate knowledge of the statistics of the errors in the model and the data. It is therefore an objective to understand in detail the sensitivity of the data assimilation scheme to the details of the defining error estimates.

APPROACH

The basic assumptions underlying data assimilation methods in use or proposed are known to be false to some degree. We plan to study the consequences of these assumptions by constructing a hierarchy of schemes with decreasing reliance on ad hoc assumptions. It is our guiding philosophy that the best way to learn how to design and implement the most economical methods that meet our needs is to begin by implementing methods which are as close to optimal as possible. From that point, we can quantify the loss of accuracy and the saving of resources associated with each simplification of the model or the data assimilation scheme.

Work is proceeding toward a theoretical basis for the next generation of data assimilation methods in which randomness and nonlinearity must be taken into account. To this end, we are applying tools from stochastic differential equations and from dynamical systems theory. Since our model systems are characterized by high dimensional state spaces, Monte Carlo methods must be used to study the behavior of the stochastic systems.

The theory of nonlinear filtering provides a framework in which problems of data assimilation with nonlinear models and non-Gaussian noise sources can be treated. In the case of linear models and Gaussian noise sources, this theory reduces to the familiar Kalman filter. In the formal theory of nonlinear filtering, the final result is not a single model state vector or trajectory in state space, but a PDF defined as a scalar function of the state variables and time. From this PDF, the mean, median, mode, or other statistic can be computed for use as the working estimate of the state of the system, along with the desired confidence intervals. The assignment of confidence limits corresponds in the case of a group of particles in physical space to drawing contours in the spatial domain which can be expected to define a region which contains, say, 90% of the particles.

The problem is that for even schematic models of the ocean or atmosphere, an unrealistically large number of particle trajectories in phase space must be calculated in order to represent the PDF faithfully. Useful ensemble analysis therefore requires judicious choice of ensemble members. We have begun a program of investigation of ensemble methods, which we see as facilitating the generation of the forecast error estimates necessary for data assimilation. These forecast error estimates are of interest in and of themselves, since they have the potential of providing a priori estimates of the reliability of a given forecast.

We are now working with a simple spectral approximation of a β-plane channel with topography (Gravel and Derome, 1993). This system was chosen as an intermediate stage between the simple problems with only qualitative resemblance to actual geophysical fluid dynamical systems, and practical working ocean models. Understanding of the behavior of the unperturbed system is a prerequisite for understanding the stochastic system, and we have used dynamical systems tools (see, e.g., Seydel, 1988) to determine the qualitative structure of the solutions for different parameter values. We have found the system to be much more complex than anyone had suspected from earlier investigation of similar systems (see, e.g., Pedlosky, 1981). In particular, we have found multiple families of limit cycles, which in at least one parameter range include multiple stable limit cycles in a region where no equilibrium point is stable. The limit cycles themselves bifurcate into invariant tori, and probably into more complex structures still. This structure is only apparent in versions of the beta plane channel with higher resolution than the models used in early qualitative studies such as Charney and DeVore (1979) or Pedlosky (1981).

This work is being done in collaboration with Dr. E. F. Carter of Taygeta Scientific, Inc. Technical support is provided by Ms. Laura Ehret. Two students, Renellys Perez and Guillaume Vernieres are now beginning work in this group, and it is anticipated that they will work directly on this project or on one with similar objectives and approach.

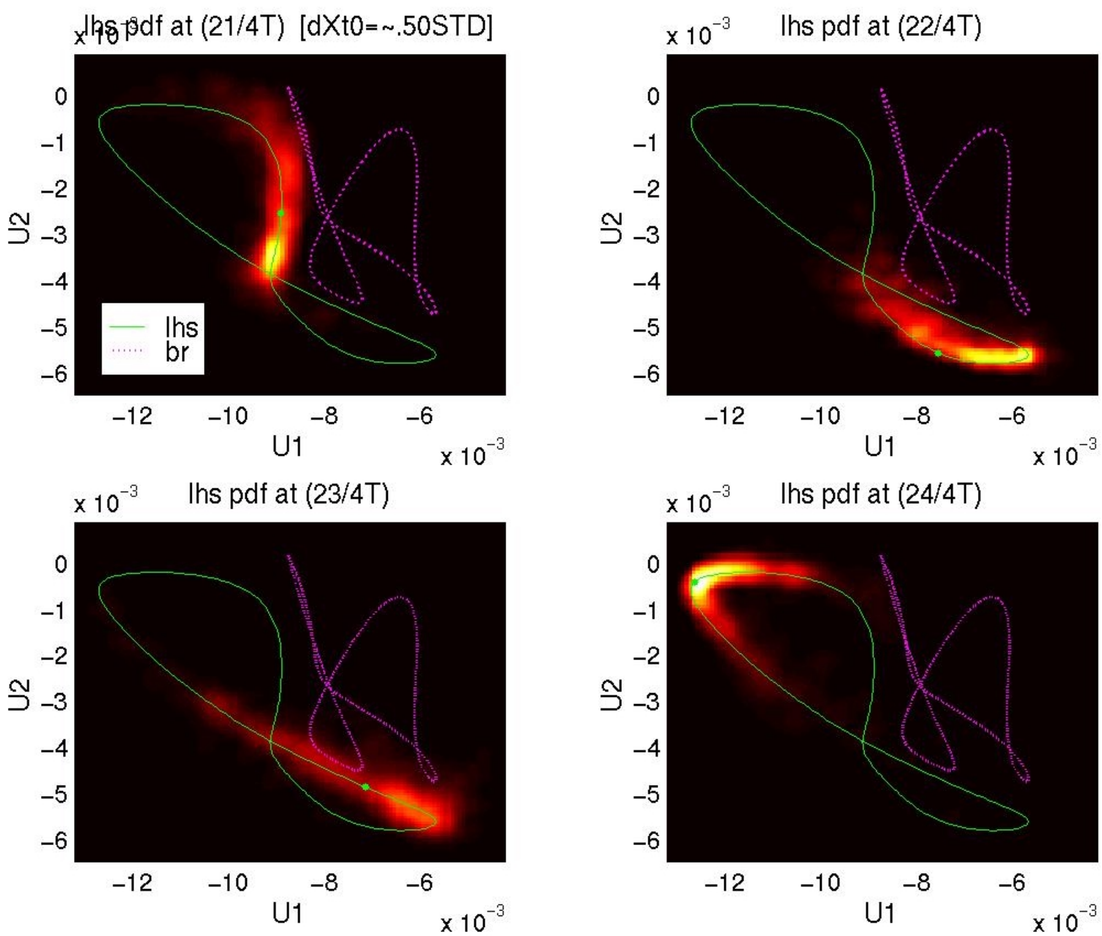
WORK COMPLETED

We have succeeded in generating breeding modes for our channel model. When we duplicate the scheme used at NCEP, we find the results to be sensitive to observation error and to the details of the underlying analysis scheme.

We have characterized the basins of attraction of the stable limit cycles of our channel model.

We have generated singular vectors of the tangent linear approximations to the co-existing stable limit cycles in our channel model. By comparing them to the results of large Monte-Carlo experiments, we have been able to evaluate the efficiency with which they can be used to characterize the evolution of the covariance field of a collection of forecasts begun with initial conditions chosen at random.

We have calculated PDF's for our channel model with random initial conditions.



Probability density function of a channel model with random initial conditions. The initial PDF is Gaussian with diagonal covariance centered on the green limit cycle. The variances are chosen so that the standard deviations are equal to one-half the component by component RMS amplitude of the limit cycle. The four frames show a detail of the evolution of the PDF during the sixth period of integration; a period here refers to the period of the green limit cycle. The time series proceeds from upper left to upper right to lower left to lower right in quarter period increments, ending at the end of the sixth period. The PDF is contoured from yellow at its maximum, decreasing to red. The other known stable limit cycle is shown in magenta. The axes are the final time singular vectors corresponding to the largest singular values.

RESULTS

For the barotropic channel model, both of the known stable periodic orbits exhibit growing singular vectors over a single period. The apparent tendency of the stochastic system to prefer transitions from one orbit to the other over transitions in the reverse direction can be attributed to the larger basin of attraction of the preferred orbit.

For the barotropic channel model in deterministic form with random initial conditions, small ensembles with tens of members chosen from singular vectors or bred growth modes do not model the covariance evolution faithfully.

IMPACT/APPLICATIONS

Major weather centers, including the US National Center for Environmental Prediction (NCEP) and the European Center for Medium-Range Weather Forecasting (ECMWF) now use ensemble methods for operational forecast validation. Our work on Monte-Carlo methods should provide enhanced capability for validation of forecasts of the ocean and atmosphere, in addition to application to data assimilation. Our work on breeding modes and planned work on other schemes for ensemble generation should provide significant guidance in optimizing methods for generation of ensembles.

TRANSITIONS

RELATED PROJECTS

“The Prediction of Wind-Driven Coastal Circulation,” a two-year program under the National Oceanographic Partnership Program (NOPP) began last summer. The objective of that two-year project is to produce a practical nowcast system for the ocean off the Oregon coast. This project includes a modeling and a field component. The NOPP team includes Professors Allen and Miller.

“Assimilation of Coastal Radar Surface Current Measurements in Shelf Circulation Models.” Work is in progress on the investigation of data assimilation systems for use with surface velocity data from coastal radar. This project is in collaboration with Professor John Allen.

REFERENCES

Charney, J.G. and Devore, J.G. 1979: Multiple flow equilibria in the atmosphere and blocking, J. Atmos. Sci., 36, 1205-1216.

Gravel, S. and Derome, J. 1993: A study of multiple equilibria in a β plane and hemispheric model of a barotropic atmosphere, Tellus, 45A, 81-98.

Miller, R.N., Carter, E.F. and Blue, S.T. 1999: Data assimilation into nonlinear stochastic models, Tellus, 51A, 167-194.

Molteni, F., Buizza, R., Palmer T.N., and Petroliagis, T. 1996: The ECMWF ensemble prediction system: Methodology and validation, Q. J. R. Meteorol. Soc., 122, 73-119.

Pedlosky, J., 1981: Resonant topographic waves in barotropic and baroclinic flows, J. Atmos. Sci., 38, 2626-2641.

Seydel, R., 1988: From equilibrium to chaos: practical bifurcation and stability analysis. Elsevier, NY. 367 pp.

Toth, Z. and Kalnay, E. 1993: Ensemble forecasting at NMC: the generation of perturbations, Bull. Am. Meteorol. Soc., 74, 2317-2330.

PUBLICATIONS

Miller, R.N., Carter, E.F., and Blue, S.T. 1998: Data assimilation into nonlinear stochastic models, Tellus, 51A, 167-194.